DOCUMENT RESUME

ED 178 578 TM 009 737

AUTHOR Lowry, Stephen R.
TITLE The Effect of Misinformation on Item Discrimination Indices and Estimation Priorities of Multiple-Choice Test Scores.
PUB DATE Apr 79
NOTE 9p.; Paper presented at the annual meeting of the American Educational Research Association (63rd, San Francisco, CA, April 8-12, 1979)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Achievement Tests; Confidence Testing; *Guessing (Tests); Higher Education; *Item Analysis; *Multiple Choice Tests; *Response Style (Tests); Scores; *Scoring Formulas; Test Items
IDENTIFIERS *Item Discrimination (Tests)

ABSTRACT

A specially designed answer format was used for three tests in a college level agriculture class of 19 students to record responses to three things about each item: (1) the student's choice of the best answer; (2) the degree of certainty with which the answer was chosen; and (3) all the answer choices which the student was certain were incorrect. Instances where the correct answer was identified as being incorrect were scored as instances of "misinformation." Other items that were not answered correctly were scored as instances of "ignorance." score was computed which was identified as being incorrect were scored as Another score was computed which was identified as "true ability." It was concluded that the additional effort required to obtain estimates of misinformation and ignorance was not justified by item discrimination indices, but was justified in providing appropriate remediation for students, depending on whether the examinee had ignorance or misinformation. (Author/CTM)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

ED178578

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

# THE EFFECT OF MISINFORMATION ON ITEM DISCRIMINATION INDICES AND ESTIMATION PROPERTIES OF MULTIPLE-CHOICE TEST SCORES

Stephen R. Lowry
University of Nebraska

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Stephen R. Lowry

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Paper presented at the Annual meeting of the American Educational Research Association, San Francisco, California April 1979 (Session D-34)

TM009 737

# The Effect of Misinformation on Item Discrimination Indices and Estimation Properties of Multiple-Choice Test Scores

Stephen R. Lowry
University of Nebraska

An evaluation of the effect of misinformation on measurement characteristics of multiple-choice tests is complicated by the following phenomenon: There is no concrete means of determining whether, and to what extent, an incorrect choice on a multiple-choice test item may be attributed to misinformation or ignorance. On a multiple-choice test it is plausible to define ignorance and misinformation according to an examinee's strategy or behavior in answering an item. When the examinee does not know the correct answer with a substantial degree of certainty, the following behavior is hypothesized. First, he eliminates the choices he believes to be incorrect with a substantial degree of certainty. Then, he quesses among the remaining choices. If the eventual choice is incorrect, this outcome may be categorized as being due to ignorance or misinformation as follows:
Ignorance. The right answer was among the choices from which the examinee guessed.
Misinformation. The right answer was one of the choices eliminated with a substantial degree of certainty.

However, because of the different levels of risk which the examinees are willing to assume as well as the difference in the extent to which the examinees follow directions, operational definitions, as defined above, may be biased. Nevertheless, it is important initially to obtain good estimates of ignorance and misinformation scores and subsequently to evaluate their relationship to examinees performance since items marked wrong due to misinformation suggest an entirely different approach to remediation than items marked wrong due to ignorance. The two-fold objective of this paper is as follows: Initially, refined estimates of misinformation and ignorance scores will be developed using a scoring procedure developed by Coombs, Milholland, and Womer (1956). The second objective is to

investigate the empirical relationship between misinformation and ignorance respectively and item discrimination indices and estimation properties of multiple-choice test scores.

Data Collection and Processing

Data for the study were collected from 3 tests in a college level agriculture class (19 students) at a large state university. The students were required to respond to multiple-choice items according to the following novel format. The format was developed using a custom printed answer sheet with three sections. In the first section, the student registered his best quess as to the correct choice on the multiple-choice items. In the second section, the student register his reason for answering each question from among the following alternatives.

1. I recognized my choice as the correct answer as soon as I saw it and didn't really consider other choices.
2. I am reasonably certain my answer was correct but I did read the other choices and give them some consideration.
3. I was able to eliminate all but one choice as incorrect and, hence, chose what I believe is the correct answer.
4. I eliminated one or more choices as definitely wrong but I am not very certain of my answer.
5. I couldn't eliminate any choices as definitely incorrect but, I had a hunch as to the correct answer.
6. I had no idea as to what the correct answer was and thus guessed at random.

In the last section, the examinee eliminated those choices he felt were definitely incorrect. The scoring procedure thus indicated that when an examinee eliminated one or more choices in section 3 and answered 1-4 in section 2, yet chose the incorrect choice in section 1, the examinee had misinformation on the item. Otherwise, the incorrect choices were the result of ignorance. The

resulting scores were consistent with other class performance by the examinees.

The data were also analyzed to evaluate the following empirical relationships.

1. Determination of the relationship between misinformation and item discrimination indices.

   In order to determine this relationship, the following statistics were necessary. For each item, the Pearson product moment correlation between item and total score and the proportion of misinformed examinees were computed. Then the proportion of misinformed examinees was correlated with the item discrimination indices.

2. Determination of the relationship of misinformation or ignorance on items to total misinformation or total ignorance respectively as well as total raw score and total true ability.

3. Determination of the effect of misinformation or ignorance as the estimation properties of multiple-choice scores.

   For each examinee, an absolute discrepancy between raw scores and true ability which is the examinees true performance level eliminating the effect of luck, the number of items on which the examinee has misinformation, and the number of items on which the examinee had ignorance were computed. Then the misinformation scores and ignorance scores were each correlated with the discrepancy between raw scores and true ability across items.

The results were as follows:

The correlations between percent misinformed and item discrimination indices were .201 (test 1), .062 (test 2), and -.513 (test 3, $P<.05$). This suggests a developing relationship as the examinees became more familiar with the scoring procedure. For test 3, the items with lower misinformation were better discriminators probably because highly discriminating items discriminate between presence and absence of information rather than between presence of information and misinformation. The results from the second analysis indicated that misinformation on items had a

significant positive correlation with total misinformation in approximately half the items for all three tests (P<.10). There also exists a significant correlation between total misinformation for test 1 and total misinformation for test 2 (P<.05) and between the same scores for test 2 and test 3 (P<.10). Thus, there exists a moderate relationship between item misinformation and total misinformation and for total misinformation across tests indicating that misinformation is predominately an attribute of the examinee rather than the item.

Misinformation on items has a significant negative correlation with raw scores and true ability score for approximately one third of the items (P<.10). Total misinformation was also negatively correlated with raw scores for all three tests (Table 1, P<.01). Ignorance on items was seldom correlated with total misinformation. It was, however, negatively correlated with true ability and raw scores for about half the items (P<.10) in tests 1 and 2. Furthermore, total ignorance was negatively correlated with true ability for all three tests (P<.01) but only negatively correlated with raw scores for test 3 (Table 1, P<.01). Thus, examinee scores are being substantially affected by ignorance and misinformation.

Table 1, as well as including the above statistics, shows the intercorrelations between total ignorance, total misinformation, raw scores, and true ability, and the discrepancy between raw score and true ability for all three tests along with their observed probability levels. Thus, misinformation is negatively correlated with the discrepancy between raw scores and true ability for 2 of the three tests (P<.10) while ignorance is negatively correlated with the discrepancy in only 1 of three tests. Thus, misinformation appears to be somewhat affecting the ability of raw scores to estimate true ability while ignorance has only a marginal affect.

Lastly, correlations of the discrepancies between raw scores and true ability were .390 between test 1 and 2 (P<.10), .509 between test 2 and 3 (P<.03) and .241 for test 2 and 3 (P>.10).

Discussion

A modest empirical relationship appears between misinformation and item discrimination indices while a moderate relationship holds between misinformation and the discrepancy between raw scores and true ability. A relationship holds between ignorance and the discrepancy between raw score and true ability only for 1 test. Thus, the additional effort required to obtain refined estimates of misinformation and ignorance is not justified in terms of substantially improved relationships with discrimination indices. The additional effort required in estimating ignorance and misinformation scores is justified for assessing ability of raw scores to estimate true ability and ultimately in providing appropriate remediation for students depending on whether the examinee had ignorance or misinformation.

Further research in a testing situation in which item discriminations are higher and examinees have higher levels of misinformation is recommended to establish a more conclusive relationship.

Table 1: Intercorrelations between total misinformation, total ignorance, raw score, true ability, and the discrepancy between raw score and true ability N=19. (Correlation/Prob > R Under Ho:ρ=0)

| | Misinformation | Ignorance | Raw Score | True Ability | \|Raw Score-True Ability\| |
|---|---|---|---|---|---|
| **Test 1** | | | | | |
| M | 1 | | | | |
| I | .058/.81 | 1 | | | |
| R | -.699/.0009 | -.334/.16 | 1 | | |
| T | -.581/.009 | -.416/.08 | .907/.0001 | 1 | |
| \|R-TA\| | -.389/.10 | .115/.64 | .394/.09 | -.030/.90 | 1 |
| **Test 2** | | | | | |
| M | 1 | | | | |
| I | .418/.07 | 1 | | | |
| R | -.835/.0001 | -.306/.20 | 1 | | |
| T | -.906/.0001 | -.534/.02 | .860/.0001 | 1 | |
| \|R-TA\| | .432/.06 | .55/.01 | -.079/.75 | -.576/.009 | 1 |
| **Test 3** | | | | | |
| M | 1 | | | | |
| I | .469/.04 | 1 | | | |
| R | -.803/.0001 | -.705/.0007 | 1 | | |
| T | -.854/.0001 | -.828/.0001 | .882/.0001 | 1 | |
| \|R-TA\| | -.123/.61 | .064/.88 | .489/.03 | .020/.93 | 1 |

## References


Burton, Nancy W. On Investigation of Item-Scoring Formulas Which Take Into Account Random Guessing, Partial Information, and Misinformation. Unpublished doctoral dissertation, University of Colorado, 1972.

Coombs, C.H., Millholland, J.E., and Womer, F.B. The Assessment of Partial Knowledge. Educational and Psychological Measurement, 1956, 16, 13-37.

Collet, Laverne. Elimination Scoring: An Evaluation. Journal of Educational Measurement. 8, 3, 209-14.

Davis, F.B. Educational Measurements and Their Interpretation. Belmont, California: Wadsworth, 1964.

Lord, F.M., and Novick, M.R. Statistical Theories of Mental Test Scores. Reading, Mass. Addison-Wesley, 1968.